Effects of missing gravity waves on stratospheric dynamics; part 1: climatology

Roland Eichinger[1,2] · Hella Garny[1,2] · Petr Šácha[3,4] · Jessica Danker[2,6] · Simone Dietmüller[2] ·
Sophie Oberländer-Hayn[5]

Abstract

Energy and momentum deposition from planetary-scale Rossby waves as well as from small-scale gravity waves (GWs) largely control stratospheric dynamics. Interactions between these different wave types, however, complicate the quantification of their individual contribution to the overall dynamical state of the middle atmosphere. In state-of-the-art general circulation models (GCMs), the majority of the GW spectrum cannot be resolved and therefore has to be parameterised. This is commonly implemented in two discrete schemes, one for GWs that originate from flow over orographic obstacles and one for all other kinds of GWs (non-orographic GWs). In this study, we attempt to gain a deeper understanding of the interactions of resolved with parameterised wave driving and of their influence on the stratospheric zonal winds and on the Brewer–Dobson circulation (BDC). For this, we set up a GCM time slice experiment with two sensitivity simulations: one without orographic GWs and one without non-orographic GWs. Our findings include an acceleration of the polar vortices, which has historically been one of the main reasons for including explicit GW parameterisations in GCMs. Further, we find inter-hemispheric differences in BDC changes when omitting GWs that can be explained by wave compensation and amplification effects. These are partly evoked through local changes in the refractive properties of the atmosphere caused by the omitted GW drag and a thereby increased planetary wave propagation. However, non-local effects on the flow can act to suppress vertical wave fluxes into the stratosphere for a very strong polar vortex. Moreover, we study mean age of stratospheric air to investigate the impact of missing GWs on tracer transport. On the basis of this analysis, we suggest that the larger ratio of planetary waves to GWs leads to enhanced horizontal mixing, which can have a large impact on stratospheric tracer distributions.

Keywords Stratosphere · Dynamics · Gravity waves · Climate modelling · Compensation · Age of air · Brewer–Dobson circulation · Polar vortex · Mixing · Transport

1 Introduction

When in the late 1970s, the resolution of global-scale weather and climate models was increased, they started to simulate excessively strong stratospheric zonal winds

✉ Roland Eichinger
  roland.eichinger@dlr.de

1 Meteorological Institute, Ludwig-Maximilians-University (LMU), Munich, Germany

2 Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany

3 Institute of Meteorology and Climatology, University of Natural Resources and Life Sciences (BOKU), Vienna, Austria

4 Faculty of Mathematics and Physics, Department of Atmospheric Physics, Charles University Prague, Prague, Czech Republic

5 Institute of Meteorology, Freie Universität Berlin, Berlin, Germany

6 Present Address: Institute for Atmospheric and Environmental Sciences, Goethe University, Frankfurt am Main, Germany

(Palmer et al. 1986; Kim et al. 2003; Alexander et al. 2010). A number of studies (Houghton 1978; Lindzen 1981; Matsuno 1892; Holton 1982, 1983) concluded that this wind bias was mainly due to the lack of a simulated drag that is generated by breaking of subgrid-scale gravity waves (GWs). Before, the error was covered by an underestimation of the poleward momentum flux, which arose from strong dissipation in the low resolution models (Palmer et al. 1986). The first practical attempt to overcome the issue was the introduction of elevated forms of orography, which enhances the generation of planetary-scale (Rossby) wave activity. In general, atmospheric waves are induced in the troposphere and propagate upward, thereby transporting momentum and energy to the middle atmosphere. As a consequence of multifaceted stability criteria (see e.g. Palmer et al. 1986; Lott and Miller 1997), the waves eventually break and dissipate their momentum and energy with extensive effects on stratospheric dynamics (e.g. Charney and Drazin 1961; Holton and Alexander 2000). The strengthened planetary waves (through elevated orography) led to enhanced eddy flux divergence and thereby reduced the biases in zonal mean wind and temperature (e.g. Wallace et al. 1983; Palmer and Mansfield 1986; Tibaldi 1986; Iwasaki and Sumi 1986). Due to the drawbacks of the unrealistic enhanced orography (e.g. excessive precipitation, see Lott and Miller 1997), however, explicit parameterisation of the vertical propagation and the breaking of orographic GWs (OGWs) became necessary. The first OGW schemes for general circulation models (GCMs) were developed as single-wave parameterisations based on two-dimensional linear stationary hydrostatic GW theory (e.g. Boer et al. 1984; Palmer et al. 1986; McFarlane 1987). One of the main tasks of these schemes was to separate the stratospheric polar night jet from the tropospheric subtropical jet by reducing the overall magnitude of the jets and creating stronger easterly wind shear in the upper troposphere (Kim et al. 2003; Alexander et al. 2010). In later OGW schemes, lower-level drag and orographic specifications were further developed and improved (e.g. Kim and Arakawa 1995; Lott and Miller 1997; Gregory et al. 1998; Scinocca and McFarlane 2000). Various versions of these parameterisation schemes are still routinely applied in GCMs for climate simulations.

Besides the aforementioned OGWs, GWs can originate from convection, frontal instabilities or from spontaneous adjustment. According to their sources, these GWs are generally named non-orographic GWs (NGWs). Due to their small spatial scales and short time scales, the NGW spectra cannot completely be resolved by climate models either and therefore have to be parameterised (see e.g. Fritts and Alexander 2003; Alexander et al. 2010). The parameterisation schemes for NGWs also base on the

pioneering work of Lindzen (1981) and Holton (1982). The development of these schemes led to a number of spectral NGW parameterisations for use in GCMs (e.g. Medvedev and Klaassen 1995; Hines 1997; Alexander and Dunkerton 1999; Warner and McIntyre 1996) that improved the representation of the middle atmosphere. NGW schemes helped for example to produce an earlier breakdown of the Southern Hemisphere winter vortex and to drive realistic stratospheric quasi-biennial and mesospheric semiannual oscillations (QBO and SAO) (e.g. Manzini and McFarlane 1998; Scaife et al. 2000; Fomichev et al. 2002; Scinocca 2003; Giorgetta et al. 2002). Furthermore, more recent studies have shown that NGW parameterisations allow realistic modelling of the stratospheric warming frequency (e.g. Richter et al. 2010) and a reasonable representation of the stratospheric transport circulation (e.g. Shepherd 2007).

The mean meridional stratospheric overturning circulation is characterised by upward motion of air in the tropics and downward motion in the middle and high latitudes. Referring to the pioneering work of Dobson et al. (1929), Brewer (1949) and Dobson (1956), this transport pattern is known as the Brewer–Dobson circulation (BDC). The BDC influences the meridional spatial distributions of trace gases such as ozone and water vapour in the stratosphere and thereby the radiative properties of the atmosphere (e.g. Solomon et al. 2010; Thompson et al. 2011; Butchart 2014). Moreover, the wave drag that drives the BDC also affects the polar vortex strength and thus dynamical downward coupling, which influences tropospheric circulation patterns (Baldwin and Dunkerton 2001; Gerber et al. 2012; Kidston et al. 2015).

Charney and Drazin (1961) have first postulated that the stratospheric overturning circulation is driven by atmospheric waves. In an attempt to quantify the influence of the particular wave types on the overall BDC driving, Butchart et al. (2010, 2011) have applied the downward control principle (developed by Haynes et al. 1991) in multi-model comparison studies. With this principle, a linear separation of the influence of planetary-scale Rossby wave driving (the Eliassen–Palm flux divergence, EPfd; Eliassen and Palm 1960) and of small-scale GW driving (GWD) is possible. The sum of the two driving mechanisms (EPfd+GWD) is then being regarded as the overall stratospheric residual-mean circulation. Butchart et al. (2010, 2011) found an approximate contribution of 70% EPfd, and 30% GWD (subdivided into 20% OGWD and 10% NGWD) to the BDC driving (tropical upwelling) at 70 hPa. However, these values vary considerably among the models, whereas the strength of the BDC is comparably equal, or in other words, models with larger EPfd show smaller GWD contribution and vice versa. This points towards a possible compensation mechanism between the different wave types. However, since there are too many

differences between the analysed models (resolution, physics, chemistry etc.), clear conclusions about this effect could not be drawn in the studies by Butchart et al. (2010, 2011). Thereafter, Cohen et al. (2013) investigated the compensation mechanism in detail using an idealised GCM and discovered that perturbed forcings in parameterised wave driving are often canceled by resolved wave driving of opposite sign. Sigmond and Shepherd (2014) confirmed this result in a study with a comprehensive climate model and found that OGWD changes disturb the zonal winds in the upper flanks of the subtropical jet, which, in consequence, changes the wave refraction properties and hence the vertical propagation of waves. In continuing work on the topic, Cohen et al. (2014) identified three particular mechanisms that influence the interaction between resolved and unresolved drag. The first one is a stability constraint that is applied when GWs drive the stratospheric flow to go unstable and it is expected to apply mainly outside of the surf zone or in the upper stratosphere and mesosphere. The second mechanism is associated with mixing and bases on the fact that large-scale Rossby waves mix potential vorticity (PV). Planetary waves flatten the PV surfaces in the surf zone and mix away PV change induced by GWs. The third mechanism is a modification of PW propagation through non-local refractive index changes. This is most likely for GW perturbations near, but outside, the region of PW breaking for broad and weak perturbations. On the basis of these findings, Cohen et al. (2014) proposed a modified (PV-based) approach to investigate the relative roles of Rossby and GW driving on the BDC.

In the present study, our aim is to extend our knowledge of the above described interactions between the different atmospheric wave types and their influence on stratospheric dynamics. For this, we apply a state-of-the-art GCM (the EMAC model: ECHAM MESSy Atmospheric Chemistry, Jöckel et al. 2005, 2010, 2016) and analyse the impact of a simple deactivation of the OGW and of the NGW parameterisation scheme, respectively. As progression to the studies by Cohen et al. (2013) and Sigmond and Shepherd (2014), we here use a high-top full GCM and analyse both hemispheres for both NGW and OGW perturbations. This enables us to analyse the processes more comprehensively than in the aforementioned studies. In addition to that, we also study the age of air (AoA) tracer, which allows us to determine the influence of dynamical changes on stratospheric tracer transport. In Sect. 2, we describe the model setup and our analysis methods. Sect. 3 contains the results of the stratospheric model response to missing GWD by means of various dynamical variables and AoA. In Sect. 4, we discuss a few remarkable results in a broader context before we present our concluding remarks in Sect. 5.

2 Model data and methods

2.1 Setup and simulations

We apply the EMAC (ECHAM MESSy Atmospheric Chemistry, v2.53.0, Jöckel et al. 2010, 2016) model in a T42 horizontal ($\sim 2.8° \times 2.8°$) resolution with 90 layers in the vertical and explicitly resolved middle atmosphere dynamics (T42L90MA). In this setup, the uppermost model layer is centred at around 0.01 hPa and the vertical resolution in the upper troposphere lower stratosphere region (UTLS) is 500–600 m. The time step of all simulations performed for this study is 540 s and the data output interval was set to 6 h. This output interval was chosen to retrieve all model data in consistence with the dynamic variables (temperature and 3-D winds). For consistent calculation of the residual circulation ($\overline{v}^*$, $\overline{w}^*$) as well as of the horizontal and vertical Eliassen–palm flux divergence from the Transformed Eulerian Mean (TEM) equations (Andrews 1986; Andrews et al. 1987), these should have a temporal resolution that provides continuous data of the same time of the day.

We conducted three so-called time slice simulations. For these simulations, the boundary conditions [radiatively active greenhouse gases (GHGs) as well as sea surface temperatures (SSTs) and sea-ice concentrations (SICs)] of a certain climate state are periodically repeated for every simulation year to achieve a climatological mean of the climate state and an estimate of its internal variability. We realised a year 2000 climate state by using the monthly mean GHG (CO_2, CH_4, N_2O and O_3) fields averaged over the period 1995–2004 from the EMAC RC2-base04 model simulation, which was performed within the ESCiMo project (Earth System Chemistry integrated Modelling, see Jöckel et al. 2016). Note that this prescription of the radiatively active gases can result in GHG distributions that are not reflecting the resulting BDC. The three simulations were integrated over 30 years, whereas the first 10 years are considered as spin-up period to achieve steady-state conditions. Hence, only the 20 years after the spin-up are analysed. As in the ESCiMo RC2-base04 simulation, we used the SST and SIC data fields of the Hadley Centre Global Environment Model version 2— Earth System (HadGEM2-ES) Model (Collins et al. 2011; Martin et al. 2011).

The flexible structure of the Modular Earth Submodel System (MESSy, Jöckel et al. 2005) allows us to use the same executable in all three time slice simulation, the differences between them are realised through changes in so-called namelist settings (see Jöckel et al. 2005). This ensures that no numerically-caused or machine-dependent differences can occur between the simulations. In the standard reference (REF) setup, we use the basic EMAC modules for dynamics, radiation, clouds and diagnostics

(namely, AEROPT, CLOUD, CLOUDOPT, CONVECT, CVTRANS, E5VDIFF, ORBIT, OROGW, PTRAC, RAD, SURFACE, TNUDGE, TROPOP, VAXTRA, refer to Jöckel et al. 2005, 2010, 2016, for details on these submodels). In these submodels, the default EMAC settings are used (see Jöckel et al. 2005, 2010, 2016). No (interactive) chemistry is used in the simulations. We conducted the REF simulation and two sensitivity simulations. The setups of the latter two simulations differ in a) the deactivation of the module OROGW, which accounts for orographic GW forcing (noOGW) and b) the deactivation of the module GWAVE, which accounts for non-orographic GW forcing (noNGW). These changes result in the omission of the tendency addition (see e.g. Eichinger and Jöckel 2014) to temperature and 3-D winds due to the process/module in question. Thereby the respective GW scheme does not influence model dynamics. Note that a simulation with both GW schemes switched off is not feasible in the same manner. Attempts have shown that this causes extremely fast zonal wind speeds so that the Courant–Friedrichs–Lewy criterion is violated, which makes the model unstable.

We use the default EMAC GW drag schemes in our simulations. The non-orographic GW module GWAVE (Baumgaertner et al. 2013) was originally developed by Hines (1997). The launch level where GWs are released is set to be near 643 hPa and the namelist parameter *rmscon*, which controls the momentum deposition in the stratosphere and mesosphere, is set to 0.92 to achieve an optimal strength of the Antarctic polar vortex (see Jöckel et al. 2016, for more details). Orographic GWs are parameterised with the columnar approach by Lott and Miller (1997) in the module OROGW.

2.2 Analysis methods

In this study, we apply a number of diagnostics to analyse the results of the model simulations. The principles as well as some information about the application of these analysis methods are given in the following.

For the description of dynamical processes in the stratosphere, we use the traditional TEM equations (Andrews 1986; Andrews et al. 1987) and the dynamics of planetary waves are diagnosed by the the Eliassen–Palm–flux vector ($F = (0, F^{(\phi)}, F^{(z)})$, see Andrews et al. 1987). The latter is directly proportional to the meridional eddy heat and momentum flux (plus ageostrophic components) through waves. Hence, the divergence of the Eliassen–Palm (EP) flux shows sources and sinks of wave energy. In the quasi-geostrophic (QG) limit, the EP-flux is proportional to the potential vorticity flux (Edmon et al. 1980). Our calculation of the EP-flux (the propagation of resolved waves) follows the method of Edmon et al. (1980) and Andrews et al. (1983) for logarithmic pressure coordinates. Note that in

our analyses the scale of the arrows differs in all the panels and is not denoted. Hence, the depiction of the EP-fluxes does not serve to show arrows of the correct magnitude, but only to show regional differences within one panel and their directions (see also Andrews et al. 1983). Since the majority of the GW spectrum is not resolved by the GCM, this diagnostic is dominated by Rossby waves (resolved). Therefore, we refer to the EPfd as the resolved wave drag. The GWD (= OGWD + NGWD) complements the latter to the full stratospheric mean flow forcing and is a direct output from the two above mentioned parameterisation modules OROGW and GWAVE.

Similar to the study by Cohen et al. (2014), we calculate the quasi-geostrophic potential vorticity (q) gradient to apply PV-based diagnostics of mixing. In the quasi-geostrophic limit, the zonal-mean meridional PV flux is equal to the resolved wave driving. Cohen et al. (2014) have shown that this quantity can provide insights into location and strength of the mixing barriers and the well-mixed surf zone.

Matsuno (1970) derived a two-dimensional wave equation to define the refractive index (RI) n^2. Wavelike solutions are only possible when n^2 is positive, which means that waves can only propagate through regimes with positive n^2. We follow the n^2 formulation by Gerber (2012):

$$n^2 = \frac{\overline{q}_\Phi}{r_0 \cdot (\overline{u} - c)} - \left(\frac{k}{r_0 \cdot cos\Phi} \right)^2 - f^2 \cdot F(N^2). \qquad (1)$$

Here, the overbars represent a zonal mean, c is the phase velocity (which we set to 0 assuming stationary waves), k is the wave number, $\overline{q}_\Phi / r_0 = \overline{q}_y$ is the meridional PV gradient, u the zonal wind, Φ the latitude and $f = 2 \cdot \Omega \cdot sin\Phi$ the Coriolis parameter. We follow Gerber (2012) and Simpson et al. (2009) approximating N (the Brunt-Väisälä frequency) to be constant, which reduces the function $F(N^2)$ to $1/(2NH)^2$, where H is the scale height. According to Gerber (2012), this does not lead to qualitative differences in the results (for details see also Harnik and Lindzen 2001). Since the second term on the right side of Eq. 1 simply produces a wave number-dependent offset and we mostly focus on differences in the present study (for which case the term is canceled out), we neglect this term and thus do not have to specialise on a particular wave number.

We diagnose the large scale advective transport using the mass-flux streamfunction, which is calculated by means of two different methods: (a) the "direct" method via the residual circulation (in the following indicated by "DIR"; using $\overline{v}^*$ as in Oort and Yienger 1996) and (b) the downward control (DWC) method following Haynes et al. (1991). The latter principal is used to separate the forcing of the mass flux into the three different parts, EPfd, OGWD and NGWD. The streamfunction of the residual circulation $\overline{\chi}^*$ can then be described by

$$\overline{\chi}^*_{\overline{m}_0} = \int\limits_p^0 \left[\frac{1}{cos\phi \cdot f} \left(\frac{1}{r \cdot cos\phi} \nabla \cdot F - \frac{\partial \overline{u}}{\partial t} + \overline{X} \right) \right]_{\phi = \phi(\overline{m}_0)}, \tag{2}$$

with $\nabla \cdot F$ denoting the EPfd, $\overline{u}$ the mean background meridional wind, $\overline{X}$ the total zonal GW drag (OGWD + NGWD), f the coriolis parameter, ϕ the given latitude, p the pressure and $\overline{m}_0 = r \cdot cos\phi(\overline{u} + r \cdot \Omega cos\phi)$ the climatological angular momentum on the contours of which the integration takes place. For steady-state conditions, which approximately are given in the winter and summer seasons, the sum of all three forcings, in theory, is equal to the direct method (see e.g. Randel et al. 2008; Sato and Hirano 2019) and is in the following referred to as ALL. Since the downward control method cannot be applied in the tropics due to a division by the Coriolis parameter, the results are only shown poleward of 20°S and 20°N.

For qualitative analysis of wave compensation and amplification, we use the compensation index. This index was introduced by Cohen et al. (2013) as a heuristic measure of the compensation strength. It is derived by defining a perturbation and a response to a system which has reached a steady-state. Here, the system in equilibrium is the reference simulation and the perturbation is the respectively missing GW forcing of the particular sensitivity simulation. The response is hence the newly adjusted equilibrium of the sensitivity simulation. Mathematically, Cohen et al. (2013) describe the compensation index as

$$C = -2 \cdot \frac{\sum_i \left[P(x_i) \cdot R(x_i) \right]}{\sum_i P^2(x_i) + \sum_i R^2(x_i)}, \tag{3}$$

where P is the perturbation to the zonal-mean wave driving and R the response, or the net change in the other components of the zonal wave driving. x_i is the spatial coordinate. For our application, this means that for example for the case of the comparison of the reference simulation with the simulation without orographic GWs

$$C_{noOGW} = -2 \cdot \frac{\Delta\overline{G}_{OGW} \cdot \left(\Delta\overline{G}_{EPfd} + \Delta\overline{G}_{NGW} \right)}{\Delta\overline{G}_{OGW}^2 + \left(\Delta\overline{G}_{EPfd} + \Delta\overline{G}_{NGW} \right)^2}, \tag{4}$$

where

$$\Delta\overline{G}_j = j_{noOGW} - j_{REF} \tag{5}$$

and j stands for the three different driving forces OGW, NGW and EPfd. To show the zonal mean of C in all latitudes and altitudes, we do not sum over the spatial coordinates x_i as Cohen et al. (2013). $C = 1$ means perfect compensation of the missing GW drag through the other wave driving components and $C = -1$ signifies that the other wave driving

components amplify the effect of the missing GW drag on the overall streamfunction.

Stratospheric mean age of air (AoA) is defined as the mean residence time of an air parcel in the stratosphere (Hall and Plumb 1994; Waugh and Hall 2002). In EMAC, the AoA tracer is implemented as an inert tracer with a mixing ratio that linearly increases over time as a global lower boundary condition ("clock tracer", Hall and Plumb 1994). AoA is then calculated as the time lag between the local mixing ratio at a certain grid point and the current mixing ratio at a reference point. To compute AoA consistently with the residual circulation transit times (RCTT, see below), we subtract the average of the zonal mean thermal tropopause AoA value in the tropics (20°S–20°N) from the actual AoA value at each grid point. The RCTT is the time that air would need from the tropopause to a given point in the stratosphere, if it only followed the residual circulation (without any mixing or diffusion processes). RCTTs are calculated via backward trajectories on the basis of the TEM meridional ($\overline{v}^*$) and vertical ($\overline{w}^*$) velocities (referred to as residual velocities). The RCTT is then the time that these backward trajectories require to reach the tropopause from the respective starting point in the stratosphere. The difference between AoA and RCTT is defined as aging by mixing (A_mix; Garny et al. 2014). This includes all remaining physical and numerical processes, including mixing on resolved and on unresolved scales. For more details see also Birner and Bönisch (2011), Garny et al. (2014) or Dietmüller et al. (2017). As a diagnostic of the relative strength of mixing, we calculate the mixing efficiency ϵ, which is defined as the ratio of the mixing mass flux to the net mass flux between the tropics and the extratropics. Details on the calculation of the mixing efficiency are given in the supplement.

3 Response to absence of GWs

3.1 Dynamics and circulation

It was already detailed out in the introduction that, historically, the main intention of introducing GW parameterisations to GCMs was the correct representation of the zonal wind in the stratosphere, in particular, the subtropical jets and the polar vortices. Thus, Fig. 1 shows the zonal mean zonal wind ($\overline{u}$) differences between the REF simulation and the two sensitivity simulations (noNGW and noOGW), for DJF and JJA.

The polar vortices strongly accelerate when GWs are missing. This is what we expected from historical model simulations without GWs (see Sect. 1, Palmer et al. 1986; Alexander et al. 2010). One of the first issues that were tackled with GW parameterisation schemes was to reduce the excessively strong modeled winds in the mid to high
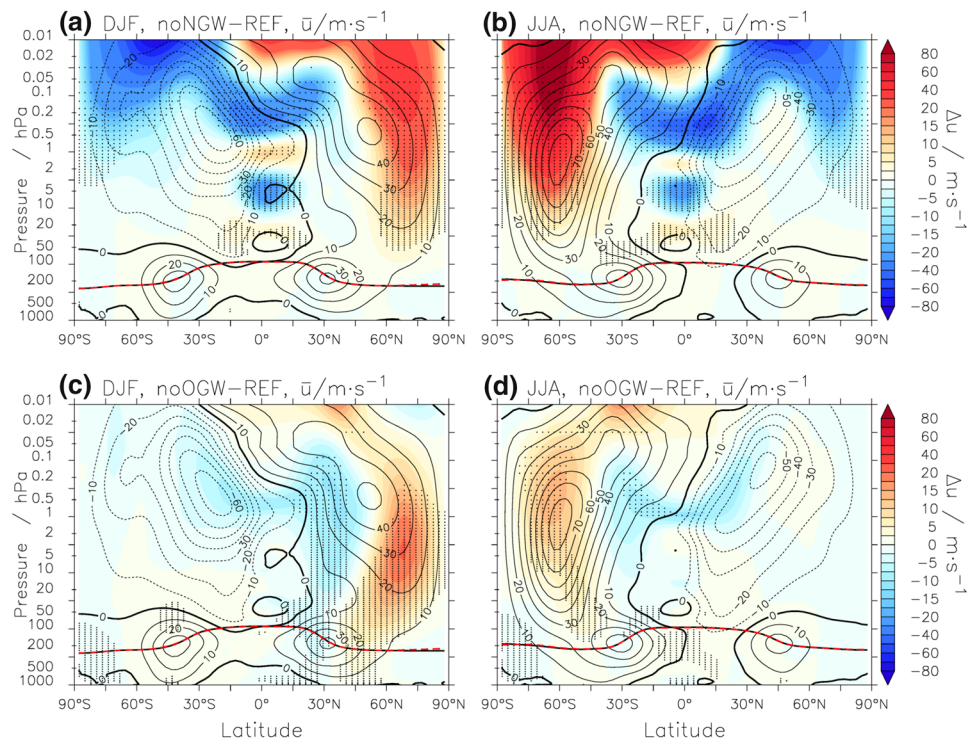
latitudes, i.e. the polar vortices (see e.g. Kim et al. 2003). In the noNGW case (Fig. 1a, b), the maxima of the polar vortices (and the easterly winds in the respective summer hemisphere) shift upward and poleward. In the tropics, a pattern can be seen that alternates with altitude. This could be expected as well, because the model has an internally generated QBO (Giorgetta et al. 2002), which strongly depends on the forcing of NGWs (as it has been analysed e.g. by Scaife et al. 2000; Giorgetta et al. 2002). The simulation without NGWs is stuck in a QBO phase with easerly winds in the upper stratosphere and westerly winds in the lower stratosphere (not shown).

Figure 1c and d show that the response due to the absence of OGWs is generally weaker compared to the response to the absence of NGWs. The polar vortices shift downward and the subtropical jet streams poleward. This means that the valve layer between the jets weakens and the jets tend to merge. As mentioned above, OGW schemes initially were applied to separate the stratospheric polar night jet from the subtropical jet by creating stronger easterly wind shear in the upper troposphere. OGW parameterisations have a large forcing on the mid latitude lower stratosphere, which has a direct impact on the subtropical jet streams (Kim et al. 2003). Moreover, in the noOGW case, some significant differences can be seen at the surface. This can be a downward propagating signal from the stratospheric wind anomalies, or an effect caused by blocking of the low level flow due to sub-grid scale orography. The latter process is also included in the here applied OGW scheme by Lott and Miller (1997)

and it can alter the zonal mean winds at the surface. Hence, we cannot disentangle how much of the changes at the surface are due to low-level flow blocking or due to a downward propagating signal from the stratosphere.

Modified patterns of the zonal mean winds alter the atmospheric conditions of wave propagation and refraction. Sigmond and Shepherd (2014) already showed that a change in GWD leads to a modification of the vertical propagation of resolved waves. To analyse how the atmospheric conditions for wave propagation change in our simulations, Fig. 2 shows the refractive index (RI) climatologies of the REF simulation and the RI differences of the sensitivity simulations to the REF simulation in DJF and JJA, respectively.

The RI (Fig. 2) shows positive values in the mid to high latitudes of the winter hemispheres, where waves can propagate vertically and meridionally. The maxima can be seen along a transition region to the tropics and in the mid latitude lower stratosphere. In DJF (Fig. 2a), the latter region reaches to the pole, in JJA (Fig. 2d) the maxima spread out to higher altitudes. The patterns of the individual model setups generally look similar (the RI climatologies of the sensitivity simulations are shown in Fig. S1 in the supplement), but there are a few differences between the REF and the sensitivity simulations. In DJF in both sensitivity simulations (Fig. 2b, c), the RI is reduced in large parts of the mid latitudes and mostly enhanced in the upper stratosphere and mesosphere. In JJA (Fig. 2e, f), the RI changes are weak in the noOGW case, but in the noNGW case, strong changes can be seen from the mid stratosphere up to the mesosphere. There is a
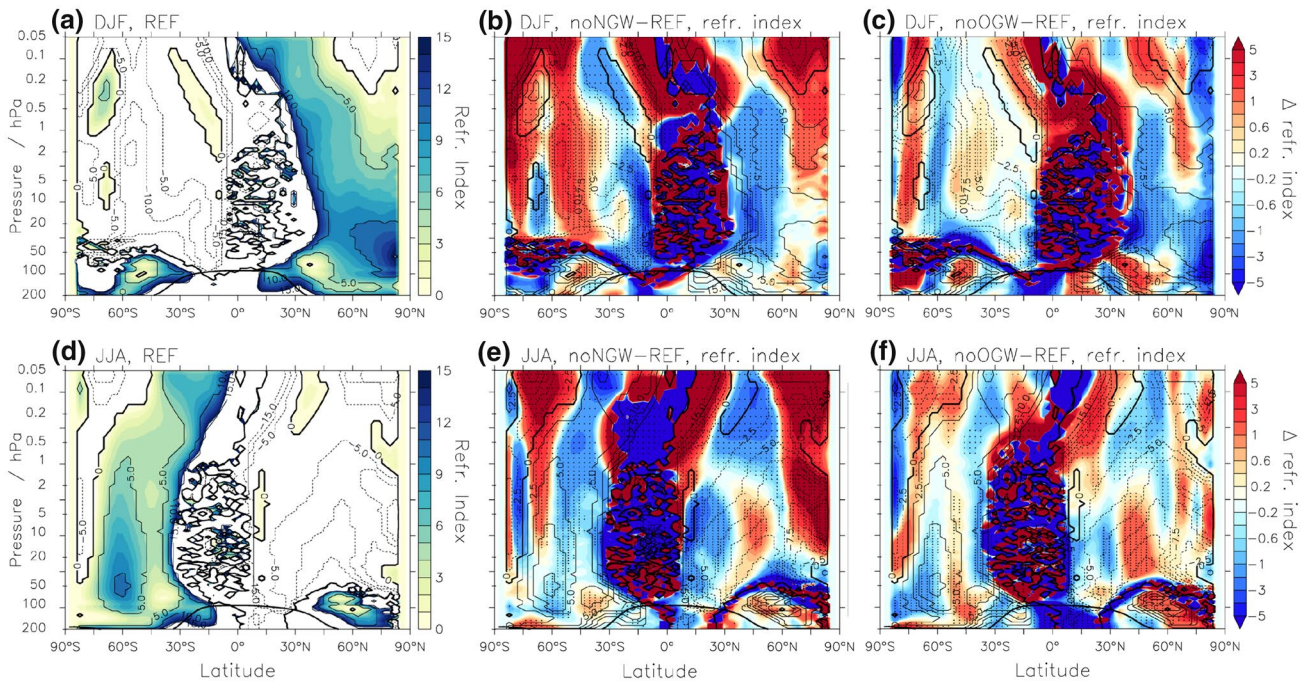
Fig. 2 Refractive index climatologies of the REF simulation (**a**, **d**) and differences of the noNGW (**b**, **e**) and noOGW (**c**, **f**) simulations to the REF simulation for DJF (**a**, **b**, **c**) and JJA (**d**, **e**, **f**). In the climatologies (**a**, **d**), negative values are not shown because a negative RI means that no wave propagation can take place. Note that the colour bar in a and d is cut above an RI of 15, because in the tropics the index is blurred due to the calculation method. In the differences (**b**, **c**, **e**, **f**), the contour lines show the refractive index climatology of the REF simulation of the respective season and the thick black line denotes the corresponding tropopause. The dotted regions mark where the differences are significant on the 95% level

stripe of enhanced RI values in the high latitudes and a stripe of reduced values in the mid latitudes. These differences are most pronounced in the mesosphere.

These changes of the wave refraction properties have a large influence on the vertical propagation of resolved waves and thereby on the strength and pattern of momentum and energy deposition. Figure 3 shows the EPfd and EP-flux vector climatologies of the REF simulations and the differences between the REF and the noNGW/noOGW simulations. The climatological patterns of EPfd and EP-fluxes of the sensitivity simulations are provided in the supplement (Fig. S2).

In the DJF differences (Fig. 3b, c), enhanced vertical propagation of resolved waves (EP-fluxes) can be seen in the high latitudes for both sensitivity simulations in the stratosphere. The enhanced vertical propagation causes less drag in the lower stratosphere and more drag in the upper stratosphere. In the noOGW case (Fig. 3c), the equatorward propagation of the resolved waves is enhanced in the low to mid latitudes. Also, poleward of 60°N up to about 10 hPa the upward propagation of PWs is slightly suppressed. There is a patch of reduced EPfd in the lower stratosphere mid latitudes (the region of the climatological OGWD maximum, see Fig. S2), which is flanked by two patches of enhanced EPfd. This can be attributed to a missing impact of OGWD above the subtropical jet,

because similar anomalies of the EPfd have been documented in idealised model studies by Šácha et al. (2016) and Samtleben et al. (2019). In those studies, artificially injected GWD in this region inhibits the upward and equatorward PW propagation from the mid latitudes and focuses the PWs poleward creating an EPfd anomaly in the polar region. The EPfd anomaly on the equatorward flank is also present in the idealised studies, it may be connected with in situ creation of PWs. These changes are in agreement with the zonal wind deceleration equatorward from the center of the subtropical jet (Fig. 1c). In the absence of OGWD, PWs can propagate and break here. As a consequence, PWs inside the polar vortex are reduced, and the winds poleward of the subtropical jet maximum accelerate. The EPfd anomaly in the high latitude upper stratosphere and lower mesosphere is probably connected with the enhanced zonal winds in the polar vortex that support PW propagation. The initially smaller EP-fluxes converge in higher altitudes without the presence of OGWD.

In the noNGW case in DJF (Fig. 3b), the equatorward and upward directed EP-flux vectors are enhanced only above 50 hPa. In JJA (Fig. 3e), the EPfd changes look similar to those in DJF, but the decrease in the high latitude lower stratosphere is much weaker. The EP-fluxes in JJA in the noNGW case (Fig 3e), however, are reduced in most parts of
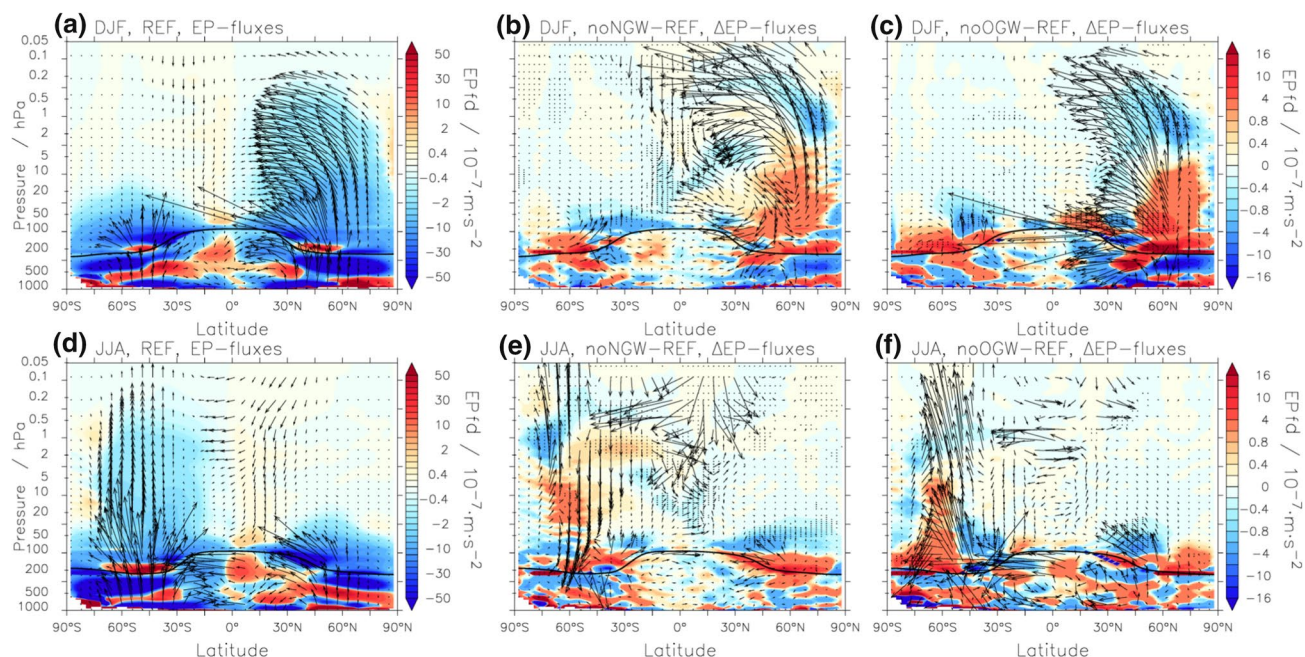
Fig. 3 EPfd (colours) and EP-flux vector (arrows) REF climatologies (**a**, **d**) and differences between the noNGW and REF (**b**, **e**) and the noOGW and REF (**c**, **f**) simulations for DJF (**a**, **c**) and JJA (**b**, **d**).

The thick black line shows the REF tropopause and the dotted regions mark where the EPfd differences are significant on the 95% level. In the troposphere, all values have been scaled by 0.1 for visual purposes

the stratosphere, only in the high latitude upper stratosphere they are enhanced. This behaviour is in contrast to all other investigated cases.

Hence, the sensitivity simulations not only differ by the lack of the respective GWD, but also by restructured EPfd patterns. As the meridional overturning circulation is driven by atmospheric waves (Charney and Drazin 1961), these wave forcing changes have an impact on the BDC. To analyse this, Fig. 4 displays the differences in zonal mean meridional mass streamfunction ($\overline{\chi}$, as calculated with the "direct method" via $\overline{v}^*$) between the noNGW/noOGW simulation and the REF simulation.

Without NGWs, the strengths of the shallow and of the deep BDC branch decrease in DJF (Fig. 4a), but only small parts of that change are significant. In the upper stratosphere and mesosphere, the changes are, albeit small, widely significant. In these altitudes, we must consider that due to the low density, the absolute mass flux is low and therefore the relative changes are still large and cause the significance. NGWs help to simulate a realistic mesospheric SAO. In particular, NGWs increase the strength and the variability of the SAO and help to generate an eastward phase (see Scaife et al. 2000; Giorgetta et al. 2002). This can lead to the significant differences in the climatologies of the streamfunction between the REF and the noNGW simulations. In the tropics, the anomalies are influenced by the missing QBO in the noNGWD simulation. The QBO phase affects the net

tropical upwelling (e.g. Flury et al. 2013). Due to the stuck QBO phase in noNGW, the mass flux entering the stratosphere is lower (see Fig. S3), which can lead to a general reduction of the BDC in the noNGW simulation. In JJA (Fig. 4b), the changes without OGWs are stronger in both BDC branches and most of these changes are significant. The impact of not including OGWs (Fig. 4c, d) is a weakening of the shallow branch and a strengthening of the upper part of the deep branch, however, the latter is mostly not significant. The stronger deep BDC branch here can possibly be explained by the fact that the suptropical jet maximum shifts upward and polward and merges with the polar night jet (see Fig. 1c). Due to enhanced EPfd, the residual circulation above the shallow branch thereby accelerates. The shallow BDC branch is particularly sensitive to the lack of OGWs, it decelerates on both hemispheres but only during the winter is this change significant.

Without the presence of OGWs, enhanced vertical EP-flux leads to more EPfd in the extratropical lower stratosphere (see Fig. 3c). In the region where OGWD usually shows a maximum (see Fig. S2.2 in the supplement), resolved waves can dissipate in the noOGW simulation. In the upper stratosphere and mesosphere, resolved wave dissipation increases. In the noNGW simulation, the shallow BDC branch changes are mostly not significant in DJF (Fig. 4a). The NGW forcing is not particularly strong in the region of the shallow branch, but in JJA, a significant

shallow branch deceleration can be observed (Fig. 4b). This can be due to the associated changes in EPfd, but cannot be concluded directly from Fig. 3 as the patterns there are too patchy. Large regions of significant deep BDC branch changes can only be observed in the noNGW case in JJA (Fig. 4b). This is the only case where the vertical EP-fluxes are reduced in the mid latitudes. In the other three cases, the vertical EP-fluxes increase in the sensitivity simulation, which leads to the compensation of the missing GWD through EPfd. In the following, we will more specifically investigate these compensation mechanisms.

Following Sigmond and Shepherd (2014) and Cohen et al. (2013), we present in Fig. 5 the residual streamfunction calculated "directly" as well as via the downward control principle (see Haynes et al. 1991, and Sect. 2.2) for separation of the individual wave forcing contributions at 10 hPa. At this altitude, the deep BDC branch is represented. This analysis allows us to study the relevance of the different wave types for the meridional circulation and possible compensation mechanisms. Here, we show the DJF and JJA climatologies of the REF simulation as well as the differences between the sensitivity simulations and the REF simulation. In Fig. S5 (see supplement), we additionally provide the corresponding figure at 70 hPa, which represents the shallow BDC branch. The climatologies of all simulations are shown in Fig. S4 for 10 hPa and in Fig. S6 at 70 hPa.

The climatologies show that the two total streamfunctions calculated with the different methods (DIR and ALL) agreee well. Differences may occur due to the fact that $\partial \overline{u}/\partial t$

from Eq. 2 is not negligible even in summer and winter (for details, see Sato and Hirano 2019), but also due to a model specific threshold for horizontal diffusion (hyperdiffusivity) which depends on the wind speed. Resolved waves (EPfd) dominate the streamfunction forcing. The NGWD is almost evenly strong across all latitudes. This is due to the crude parameterisation with globally uniform wave launching (see e.g. Jöckel et al. 2016). In the three cases in the panels b, c and f of Fig. 5 (noNGW, DJF; noOGW, DJF and noOGW, JJA), the total streamfunction changes less (in absolute values and in the respective hemisphere) than expected from the change in GW forcing (i.e. the perturbation; in c and d the OGWD differences and in b and e the NGWD differences). This is the case because EPfd changes are opposite to the direction of the perturbed GW component changes (at least in the mid latitudes), reflecting the compensation mechanism between EPfd and GWD (see Cohen et al. 2013; Sigmond and Shepherd 2014). Only for the noNGW case in JJA (Fig. 5e), the EPfd changes point in the same direction as the perturbed NGWD (around 40°S) and therefore the perturbation through missing NGWD is amplified by EPfd changes there. This leads to an additional weakening of the overall streamfunction (ALL and DIR) change compared to the forcing of NGW alone and can explain why the upper BDC branch changes significantly in this case.

The wave compensation mechanism, hence, does not take place always and everywhere. In some regions and for particular changes of the zonal mean winds, wave interaction can also be amplifying and thus cause stronger BDC changes
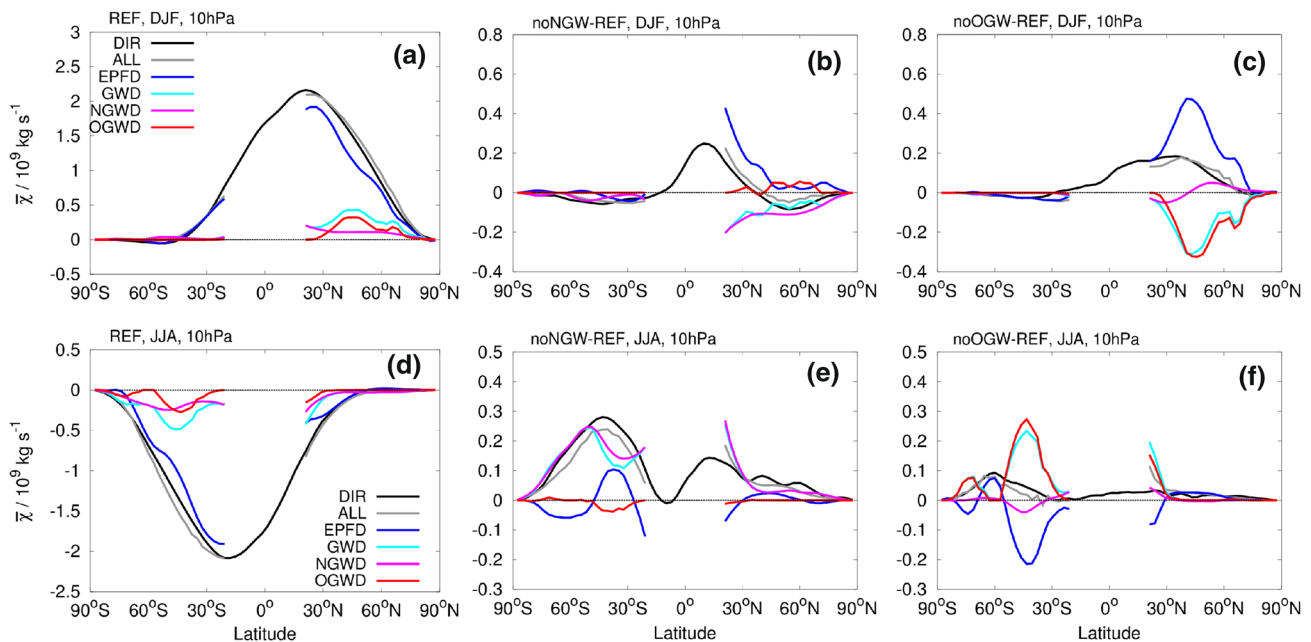
Fig. 5 Climatologies of the meridional streamfunction for the year 2000 REF simulation (**a, d**) and streamfunction differences between the noNGW (**b, e**) /noOGW (**c, f**) and the REF simulation at 10 hPa for DJF (**a, b, c**) and JJA (**d, e, f**). The streamfunction is calculated "directly" via the TEM equations (DIR, black) and with the down-ward control method for the sum (ALL, grey) of the Eliassen–Palm flux divergence (EPfd, blue) and total GW drag (GWD, light blue), which is the sum of orographic (OGW, red) and non-orographic (NGW, pink) GW drag

than the artificial perturbation alone would do. Cohen et al. (2014) discussed this mechanism to be likely for GW perturbations near, but outside the region of PW breaking, leading to non-local interactions between the perturbation and the resolved wave driving by altering the refractive properties. As we could see in Fig. 3e, the vertical EP-fluxes in JJA decrease in the noNGW case, while they increase in the other cases. Thus the resolved wave flux does not compensate the missing GWD here as in the other cases. Instead, the reduced wave flux even amplifies its effect.

Figure 6 additionally presents the compensation index of the sensitivity simulations with respect to the REF simulation for DJF and JJA. This index has been introduced as a heuristic measure by Cohen et al. (2013) and gives a qualitative view of where compensation takes place and where amplification between waves happens. The compensation index signal can be obscured by noise, if the perturbation $(P(x_i))$ is weak. Therefore, we dotted the regions where $|P(x_i)| < 0.1 \cdot max|P(x_i)|$ in the given altitude, similar as in Cohen et al. (2013). Moreover, to provide an estimation of the compensation/amplification relevance, the contour lines denote the total wave drag (EPfd + GWD) of the respective REF simulation.

In DJF (Fig. 6a, c), mostly compensation takes place. Only in the lower stratospheric high latitudes is there some wave amplification. Also in JJA in the noOGW case

(Fig. 6d), most of the wave perturbation is compensated. In the JJA noNGW simulation (Fig. 6b) in the SH, there are regions of compensation at high latitudes and in the upper stratosphere, but other regions with amplification as in the lower stratosphere and at low latitudes. Overall, this confirms the conclusion of Cohen et al. (2013) and of Sigmond and Shepherd (2014) that most parts of the changed GW forcing is compensated by the other wave driving components, but this only accounts for DJF and missing OGWs. Sigmond and Shepherd (2014) only changed the OGW forcing and analysed DJF. In JJA, however, the missing NGWD leads to large regions with wave forcing amplification in our simulations and that has a significant impact on the BDC strength. In their idealised experiments with strong and weak vortex conditions, Cohen et al. (2013) also concluded that there could be a mixture of compensation and amplification. Moreover, Cohen et al. (2014) proposed that variations in planetary wave propagation can lead to both amplification and compensation effects. Possible reasons for the amplification can be the fact that there is generally less PW activity in the SH and hence no compensation is possible or that the strong acceleration of the polar vortex distracts the vertical wave propagation. However, in Fig. 2 we also showed that the RI changes are different in the individual simulations and hemispheres and that can also cause the changes in the

Effects of missing gravity waves on stratospheric dynamics; part 1: climatology**Fig. 6** Compensation index after Cohen et al. (2013) between the noNGW (**a**, **b**)/ noOGW (**c**, **d**) and the REF simulation for the year 2000 DJF (**a**, **c**) and JJA (**b**, **d**) climatologies. The dots mark regions where the perturbation is less than 10% of the maximum perturbation of the given altitude. The contour lines show the total wave drag (EPfd + GWD) of the REF simulation and the bold black line shows the tropopause of the REF simulation. The total drag was scaled by $p/(H \cdot g)$ (where p is pressure, H the scale height 7 km and g the gravitational acceleration) to account for the mass of affected air and by $1 \cdot 10^8$ for visual purposes
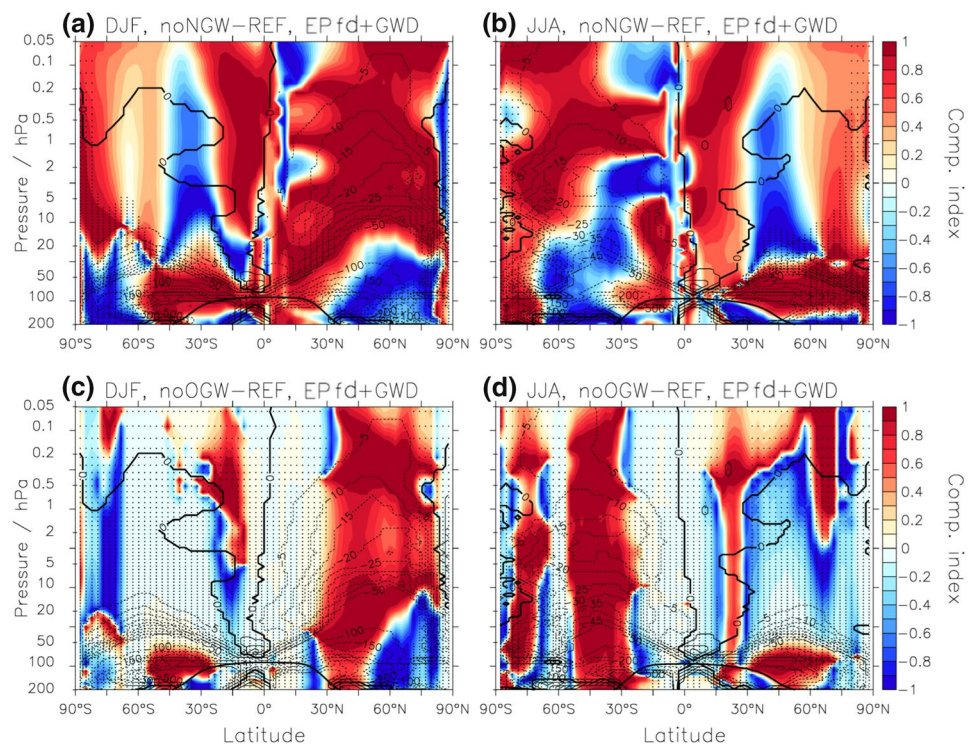
forcing. In Sect. 4, we will provide an in-depth discussion on these possible connections.

3.2 Tracer transport

Changes in BDC strength affect tracer transport and thereby the distribution of tracers in the stratosphere. Next, we therefore analyse the impacts of the dynamical changes on tracer transport changes by means of investigating AoA. As AoA is implemented as an inert tracer with a linearly rising source in the troposphere, it is a good measure for analysing changes in stratospheric tracer transport (see e.g. Hall and Plumb 1994; Waugh and Hall 2002; Garny et al. 2014; Dietmüller et al. 2017, 2018; Eichinger et al. 2019). In contrast to the streamfunction, AoA includes both, the slow overturning residual circulation transport and the effect of two-way mixing of air parcels. AoA can be separated into RCTT and A_mix (see Sect. 2.2) to display the contributions of these two processes. Figure 7 shows AoA, RCTT and A_mix differences between the noNGW/noOGW simulation and the REF simulation. Since AoA, RCTT and A_mix are integrated quantities (over several years), we provide annual means here, instead of seasonal means as for the previous diagnostics.

In the sensitivity simulations, stratospheric AoA generally increases, especially in the noNGW case and the effect is particularly strong in the lower stratosphere (Fig. 7a, d). This corresponds with the fact that the shallow BDC branch weakens in the sensitivity simulations, the consequence is

slower tracer transport. The global average AoA between 0.1 and 100 hPa is 2.86 years in the REF simulation, 3.13 years in the noNGW simulation and 3.06 years in the noOGW simulation. In the upper stratosphere, this result could have been expected in the noNGW case as well, because, as shown in the previous section, the deep BDC branch weakens there too. In the noOGW case, the deep branch is slightly enhanced, and hence AoA would be expected to decrease. However, this is not the case. In contrast, it slightly increases (Fig. 7d). The subdivision into the effects of mixing and transport can help to understand this behaviour. In the lower extratropical stratosphere, the RCTTs increase in both cases and hemispheres, reflecting the deceleration of the shallow BDC branch. However, in most other parts of the stratosphere, the RCTTs decrease in the noOGW case (Fig. 7e), in line with the deep BDC branch acceleration. In the noOGW case, the AoA increase above 20 hPa is therefore generated only through the increase of A_mix (Fig. 7f). In the noNGW case, the SH shows large areas of increasing RCTTs and in the NH, the RCTTs decrease (Fig. 7b). The SH is in line with the results from the previous section (decreasing strength of deep branch, see Fig. 4b), but the RCTT decrease in the NH was not expected. There, the streamfunction also showed a slight decrease, not an increase as the RCTTs would suggest. Again, A_mix increases strongly there, thereby overshadowing the RCTT decline and causing the AoA increase. Note, however, that when we consider changes in mixing, the most relevant regions for AoA are the mid latitudes. In-mixing of old extratropical air into the tropical pipe leads
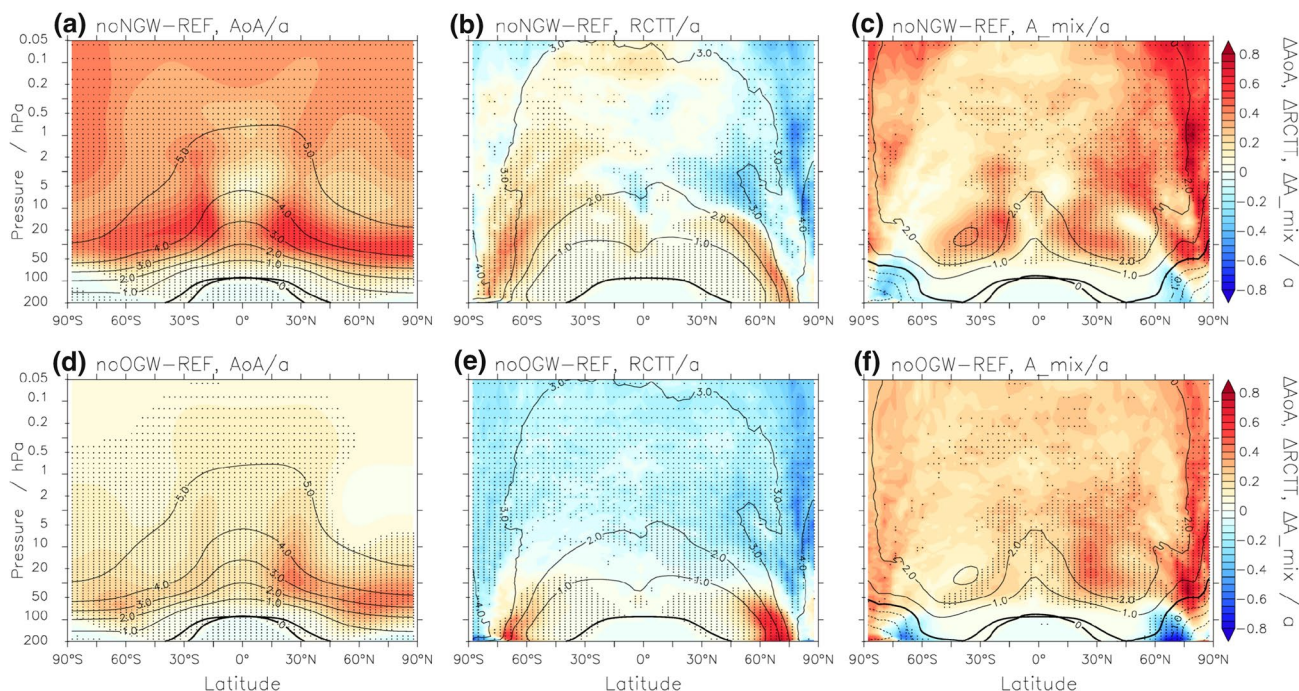
⧬ Springer

Fig. 7 Differences of annual AoA (**a**, **d**), RCTT (**b**, **e**) and A_mix (**c**, **f**) climatologies between the noNGW and REF (**a**, **b**, **c**), and the noOGW and REF (**d**, **e**, **f**) simulations, respectively. The contour lines display the climatology of the respective REF simulation. The thick black line denotes the annual mean REF tropopause and dotted regions show significance of the differences on the 95% level. The unit a stands for years (lat. anni)

to recirculation of that air around the BDC branches and thereby generates a considerable AoA increase throughout the stratosphere (see e.g. Garny et al. 2014; Eichinger et al. 2019).

To obtain insights into the reasons for the changes of mixing, we analyse the meridional PV gradient $\partial \overline{q}/\partial y$. This diagnostic provides information about meridional mixing barriers. In Fig. 8, we show the differences of the PV gradient between the two sensitivity simulations and the reference simulation for DJF and JJA.

All panels of Fig. 8 display a strong increase and a poleward shift of the PV gradient maxima in the polar vortex region. Accelerated zonal winds increase the meridional mixing barrier. The poleward shift of the mixing barrier leads to a reduction of the PV gradient in the mid latitudes, i.e. an increasing extent of the surf zone. As mentioned above, mixing in this region has a large impact on aging by mixing and thereby on AoA. Clear (and significant) differences of the PV gradient in the mid latitudes can be seen in all panels, except for the noNGW case in JJA (Fig. 8b). In Sect. 3.1, we showed that usually wave compensation takes place in the mid latitudes, which means that the missing GWD is replaced by resolved wave drag. Only in the noNGW case in JJA, wave amplification takes place, hence, resolved wave drag is reduced here too. This points towards a possible connection between the wave type ratio and the

properties of the PV gradient and thereby in horizontal mixing.

An explanation could be that the particle trajectories induced by idealized Rossby waves are, to a certain degree, ellipses with a large horizontal aspect ratio that predominantly oscillate horizontally. Therefore, we assume that the mixing induced by Rossby wave transience is also predominantly horizontal. GWs, in contrast, are not resolved in the model and their effects on mixing are not parameterized in the applied schemes. Therefore, and because GWs are small-scale waves, the breaking of GWs does not influence horizontal mixing. In the next section, we provide a detailed discussion on this hypothesis.

4 Discussion

As has been known since the 1980s, parameterisations of the drag induced by GWs are necessary to be included in stratosphere-resolving GCMs to realistically simulate the strength and pattern of the zonal winds (see e.g. Lindzen 1981; Holton 1982; Palmer et al. 1986). We conducted experiments with either the orographic or the non-orographic GW scheme switched off, thereby deliberately degrading the model climatology, in order to study the extend to which the missing

Fig. 8 Difference of the meridional PV gradient ($\partial\overline{q}/\partial y$) between the noNGW and the REF simulation (**a, b**) and between the noOGW and the REF simulation (**c, d**) for DJF (**a, c**) and JJA (**b, d**). The contour lines show the $\partial\overline{q}/\partial y$ climatology of the REF simulation of the respective season and the thick black lines denote the tropopause. Dotted regions mark where the differences are significant on the 95% level
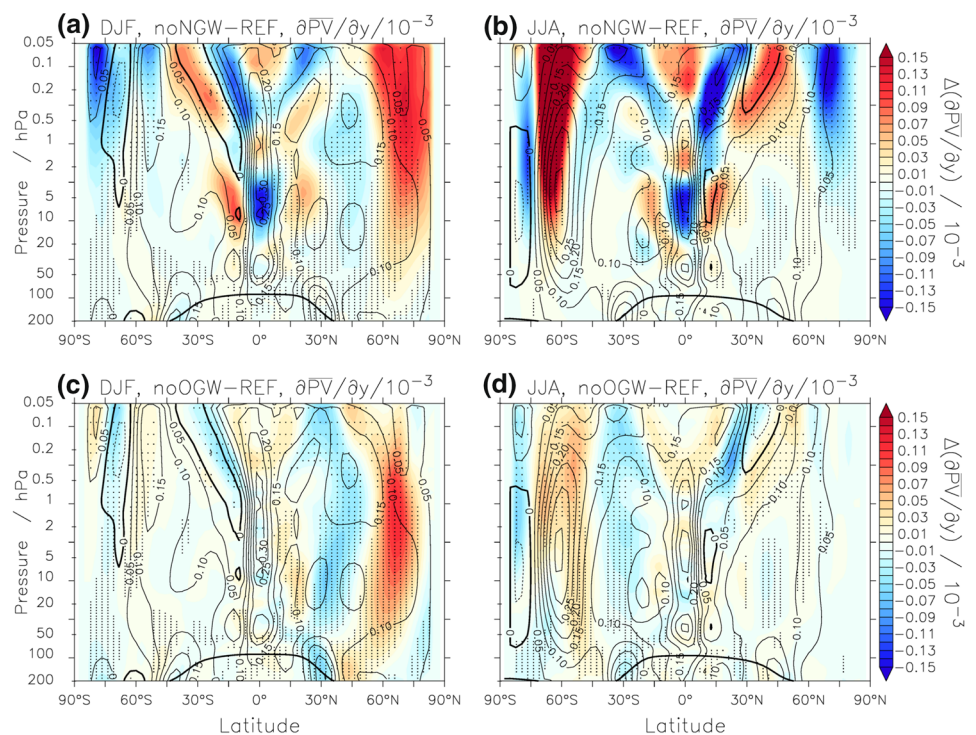
GW drag is compensated by the response of resolved waves (as suggested in recent studies).

Changes in the mean flow influence the propagation conditions for planetary waves in the atmosphere (Charney and Drazin 1961; Matsuno 1970). These are reflected in the refractive index. Therefore, the reduced GW drag (caused by the disabled GW scheme) leads to a redistribution of the total stratospheric wave drag, rather than just a reduction of it. We show that the resolved EP-fluxes within the stratosphere are altered through omission of either of the GW schemes. In most analysed cases, the wave drag redistribution results in a compensation of the missing GWD through resolved waves and thus in enhanced EPfd. This strongly reduces the net effect that a plain reduction of wave drag through missing GWs could have. Cohen et al. (2013) and Sigmond and Shepherd (2014) discussed this compensation effect, however, mostly for moderate OGW changes during DJF. Our experiments confirm their results of compensation during DJF even for complete omission of OGWs, as well as for omission of OGWs during JJA. Further, the omission of NGWs in DJF is largely compensated by resolved wave drag within the stratosphere (compensation does not happen in the mesosphere, where NGW drag accounts for almost the entire wave forcing). However, during JJA the omission of NGWs is not compensated by resolved waves. Resolved wave drag in the stratosphere is reduced in response to the omission of the NGWs, i.e. the effect of missing GW drag is even amplified. We thereby confirm the result obtained with

an idealised model by Cohen et al. (2013, 2014), namely that wave amplification can occur in response to GW forcing.

Cohen et al. (2014) suggested that compensation (or amplification) can emerge from redistribution of wave drag due to changed propagation and/or breaking conditions for planetary Rossby waves or from local Rossby wave generation due to instability in the flow. The latter is likely to occur in response to strong narrow GW torques in regions of weak resolved wave activity (i.e. outside the surf zone). Cohen et al. (2014) argue that within the surf zone, there is a constraint on the total wave force (the wave force necessary to maintain a zero PV gradient), and thus missing GW drag is compensated by enhanced propagation of planetary waves. For the omission of OGW in both winter hemispheres, we find enhanced vertical as well as meridional EP-flux propagation to the regions of "missing" GWD, suggesting that it is indeed altered propagation that causes the compensation. For the omission of NGW in DJF, vertical EP-fluxes are enhanced as well, redistributing resolved wave drag upward and compensating for the missing NGW drag there. In all three cases, the refractive index is enhanced in the upper stratosphere (the zero-line is shifted upward), as well as in the lower stratosphere (close to the climatological minimum in the refractive index), indicating that the changed propagation conditions cause the compensation. Thus, the results presented here are in line with the "local" compensation mechanism (or "PV mixing" mechanism) as suggested by Cohen et al. (2014). The zonal wind acceleration induced by the missing GWD leads to enhanced mixing barriers

(meridional PV gradients), which translate into an increase in the refractive index evoking more planetry waves to propagate into the stratosphere.

As noted above, the omission of NGWs is not compensated in the Southern Hemisphere winter (JJA). There are signs of compensation in the high latitude upper stratosphere (enhanced vertical propagation south of 60°S, in line with an enhanced refractive index) and at lower latitudes in the lower stratosphere (enhanced meridional propagation). However, reduced vertical EP-flux in the mid latitudes throughout the stratosphere additionally reduces the resolved wave forcing there. A similar change in EPfd was noted in an idealised sensitivity experiment to NGW drag by Cohen et al. (2014). They suggest that the strong change in the refractive index in the upper stratosphere at mid to low latitudes alters planetary wave propagation. The excessively strong Antarctic polar vortex may reach critical values which suppress vertical wave propagation (see Andrews et al. 1987). In our experiment, a similar decrease in the refractive index can be found. However, the vertical EP-flux is reduced already in the lowermost stratosphere (around 200 hPa, between 45–75°S), i.e. the changes in the flow below the polar night jet appear to suppress vertical wave propagation leading to a reduction in resolved wave driving in the stratosphere. The changes in the upper stratospheric refractive index are unlikely to cause the suppression of vertical EP-flux in the lower stratosphere. The refractive index has a minimum in the lowermost stratosphere (between 100 and 200 hPa), and this minimum is lower in the noNGW simulation; possibly this leads to the suppressed vertical propagation. Overall, our results suggest that the (main) reason for the amplification during JJA in response to omission of NGWs is reduced wave flux into the stratosphere, and thus there is less wave flux available to be distributed.

These changes in GWD and EPfd have a significant impact on the residual circulation. The shallow branch of the circulation generally weakens in the simulations with either of the GW schemes switched off. Particularly in the high latitude lower stratosphere, wave amplification takes place and hence, there is less total wave drag in the sensitivity simulations in the lower stratosphere. For the noNGW simulation in JJA, the missing GWD effect is amplified by resolved waves also in higher altitudes. Therefore, in the noNGW simulation, the deep branch significantly weakens in JJA. In the noOGW simulation, the deep branch rather strengthens, but this signal is mostly not significant. The change in the shape of the residual circulation (weakening of lower branch, strengthening of upper branch) with increased polar vortex strength is similar to the idealised response discussed in Gerber et al. (2012). A possible influence, however, may also originate from the fact that the QBO is stuck in one phase in the noNGW simulation. This can have an influence on the BDC strength (e.g. Flury et al. 2013).

The residual circulation changes in turn affect stratospheric tracer transport. AoA increases throughout the entire stratosphere in both sensitivity simulations, i.e. the transport circulation seems to be slowed down. The residual circulation changes account for lower transport times (RCTTs) in the SH deep branch in the noNGW case and in both sensitivity simulations in the shallow branch. Other than that, however, the residual circulation accelerates (decreased RCTTs). The reason for the AoA increase can be found in enhanced aging by mixing. In the sensitivity simulations, A_mix regionally increases strong enough to overshadow the weakening of the residual circulation. As discussed e.g. in Garny et al. (2014), Dietmüller et al. (2018) and Eichinger et al. (2019), the effect of mixing on AoA is best measured by means of the mixing efficiency, which is defined as the ratio of mixing to net (residual) mass-flux because it controls the relative increase in AoA due to mixing. The mixing efficiency ϵ (for details on ϵ, see Sect. S2.5 and e.g. Neu and Plumb 1999; Garny et al. 2014; Dietmüller et al. 2017; Eichinger et al. 2019) increases in the sensitivity simulations ($\epsilon = 0.30$ in the REF simulation, $\epsilon = 0.37$ in the noOGW simulation and $\epsilon = 0.36$ in the noNGW simulation), i.e. the relative mixing strength increases in the noNGW and in the noOGW simulations.

Through the compensation of missing GWs by resolved waves, the ratio of resolved wave driving of the BDC increases. As Rossby waves have a large horizontal aspect ratio and predominantly oscillate horizontally, we assume that, when they break, they generally lead to strong horizontal isentropic mixing. The effect of GW breaking on mixing is not explicitly parameterized in the applied GW schemes (Hines 1997; Lott and Miller 1997) and moreover, since GWs are small-scale waves, GW breaking would mix vertically rather than horizontally on global model resolutions. This can explain the coherence between the increased ratio of resolved wave driving in the sensitivity simulations and increased quasi-isentropic mixing, as reflected in aging by mixing and in the mixing efficiency. The idea of the wave type ratio influence on horizontal mixing has been pursued before without success in the multi-model study by Dietmüller et al. (2018), presumably because there were too many additional differences between the models. The results presented here indicate that the wave type ratio has an influence on the mixing strength, and thereby on stratospheric transport times. Cohen et al. (2014) state that, to change mixing, it is important where the wave perturbation takes place. The meridional PV gradient pattern changes reveal that in the noOGW case, the increase in mixing (reduced PV gradient) mostly happens in the lower stratosphere in and around the surf zone, where it has a large impact on AoA. In the noNGW case, the perturbation is more broadly distributed over the stratosphere and therefore the signal is weaker. The

PV gradient patterns also reveal that a larger ratio of PWs enhances mixing in the surf zone.

This also means that although the wave drag changes are predominantly compensating, which alleviates large parts of the residual mass circulation changes, due to its effect on mixing, the wave type ratio of the drag can still have an effect on transport. Sigmond and Shepherd (2014) concluded that the compensation effect raises the credibility of future predictions of the BDC. According to our results, however, this does not seem to hold for mixing and transport and can therefore be relevant for trace gas distributions in the stratosphere. In the second part of this study, simulations with the same setups, but for a possible future climate state will be analysed to study these effects.

Our results provide some novel insights into the interactions of different wave types with the zonal winds, the BDC and with other waves in a state-of-the-art GCM. However, in order to yield a holistic overview on the topic, we have to mention that our study bears a number of simplifications that can potentially have extensive consequences for the results and their interpretation. One of these is the columnar approach that is used in the here applied NGW scheme by Hines (1997) as well as in the OGW parameterisation by Lott and Miller (1997). This approximation is mainly used to avoid excessively high computational costs. However, for example Preusse et al. (2009) (using satellite observations) and Sato et al. (2009, 2012) (using a high-resolution GCM, which resolves the majority of the GW spectrum) have shown that GWs propagate over a considerable horizontal distance before breaking in the middle atmosphere. It is in question how this could influence the compensation effect. Also the crude parameterisation of NGW launching can lead to large errors and our zonal mean and monthly mean view on the data can hide zonally asymmetric distributions of GWD and GWD hot spots, which can have various effects on the residual circulation (Šácha et al. 2016). Geller et al. (2013) showed that GW momentum fluxes of satellite observations and GCMs are generally similar, but they also worked out shortcomings of GW parameterisations, for example a too rapid momentum flux fall off with height. Some improvements of these idealisations have recently been conducted. For example, Beres et al. (2005) and Song and Chun (2005) have included source specifications of NGWs originating from convection, Charron and Manzini (2002) and Richter et al. (2010) introduced launches of GWs from jet front systems and Amemiya and Sato (2016) developed a cost efficient OGW scheme that includes parameterised horizontal propagation. However, for various reasons, not all of these developments are included in the GW schemes that today are routinely used in GCMs, which means that also the model results that are interpreted for making predictions of the climate conditions across the 21st century (for example the models used in the CCMI-1

project, see Morgenstern et al. 2017) suffer from these inaccuracies. Therefore, the present study is crucial to better understand the processes that take place in present GCMs and to avoid misinterpretations of their results.

5 Conclusions

The known changes of the stratospheric zonal mean winds that occur when gravity waves (GWs) are reduced in GCMs lead to refractive index changes that alter the vertical propagation of waves. More specifically, the decrease in GWD leads to local acceleration of the zonal winds, which translate into increased PV gradients and thus an increase in the refractive index. In most cases, the planetary wave fluxes are enhanced by that process, which compensates the missing GW drag to some extent. We show here that compensation is active in the Northern Hemisphere winter in response to omission of either the orographic or the non-orographic GW scheme, as well as in response to the omission of orographic GWs in the Southern Hemisphere winter. However, when we omit non-orographic GWs during Southern Hemisphere winter, the EP-fluxes into the stratosphere decrease and hence, the effect of the missing GW driving is amplified. In that case, the deep BDC branch weakens strongly, while it does not show significant changes in the other cases. Thus, while we confirm the compensation mechanism via local refraction of planetary waves as suggested by Cohen et al. (2013) and Sigmond and Shepherd (2014), we argue that probably non-local effects on the flow can act to suppress vertical wave fluxes into the stratosphere for the experiment with a very strong polar vortex. In the present study, we show that this mechanism, proposed by Cohen et al. (2014), can occur for GW perturbations in a full GCM. We further showed that, when omitting parameterised GWs, the stratospheric transport circulation is affected not only by the changes in residual circulation strength, but also by changes in the relative (horizontal) mixing strength. Transport times (mean AoA) are enhanced in the simulations with missing GWs due to an increase in the relative mixing strength. We suggest that the larger ratio of planetary waves leads to stronger quasi-isentropic mixing, and thus explains the increased mixing strength. In contrast to the effect of wave drag on the residual circulation, the effect of the wave type ratio on mixing and transport cannot be compensated and hence may be important for future climate projections. However, also other factors, like the location of wave breaking, can play a role here, and will have to be analysed in specifically designed studies. Although our method includes several sources of known inaccuracies, the results of the present study are of high relevance for the interpretation of the dynamical state

of the stratosphere in climate model simulations. In the second part of this study, we will expand our analyses on the effect of missing GWs on the response of stratospheric dynamics to a future climate state.

Data availability The Modular Earth Submodel System (MESSy) is continuously developed and applied by a consortium of institutions. Use of MESSy and access to the source code is licensed to all affiliates of institutions that are members of the MESSy Consortium. Institutions can become a member of the MESSy Consortium by signing the MESSy Memorandum of Understanding. More information can be found on the MESSy Consortium website (http://www.messy-interface.org, last access: 19 December 2019). The data of the simulations can be provided by the authors upon request.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Alexander MJ, Dunkerton TJ (1999) A spectral parameterization of mean-flow forcing due to breaking gravity waves. J Atmos Sci 56(24):4167–4182. https://doi.org/10.1175/1520-0469(1999)056<4167:ASPOMF>2.0.CO;2

Alexander MJ, Geller M, McLandress C, Polavarapu S, Preusse P, Sassi F, Sato K, Eckermann S, Ern M, Hertzog A, Kawatani Y, Pulido M, Shaw T, Sigmond M, Vincent R, Watanabe S (2010) Recent developments on gravity wave effects in climate models, and the global distribution of gravity wave momentum flux from observations and models. Q J R Meteorol Soc 136:1103–1124. https://doi.org/10.1002/qj.637

Amemiya A, Sato K (2016) A new gravity wave parameterization including three-dimensional propagation. J Meteorol Soc Jpn 94(3):237–256. https://doi.org/10.2151/jmsj.2016-013

Andrews DG (1986) On the interpretation of the eliassen-palm flux divergence. Q J R Meteorol Soc 113:323–338. https://doi.org/10.1002/qj.49711347518

Andrews DG, Mahlman JD, Sinclair RW (1983) Eliassen-palm diagnostics of wave-mean flow interaction in the gfdl "skyhi" general circulation model. J Atmos Sci 40:2768–2784. https://doi.org/10.1175/1520-0469(1983)040<2768:ETWATM>2.0.CO;2

Andrews DG, Holton JR, Leovy CB (1987) Middle atmosphere dynamics, vol 40. International geophysics. Elsevier Science. ISBN: 9780120585762, 9780080511672

Baldwin MP, Dunkerton TJ (2001) Stratospheric harbingers of anomalous weather regimes. Science 294:581–584. https://doi.org/10.1126/science.1063315

Baumgaertner AJG, Jöckel P, Aylward AD, Harris MJ (2013) Simulation of particle precipitation effects on the atmosphere with the MESSy model system. In: Lübken F-J (ed) Climate and weather of the sun-earth system (CAWSES). Springer Atmospheric Sciences, Springer, Netherlands, pp 301–316. https://doi.org/10.1007/978-94-007-4348-9_17

Beres JH, Garcia RR, Boville BA, Sassi F (2005) Implementation of a gravity wave source spectrum parameterization dependent on the properties of convection in the whole atmosphere community climate model (waccm). J Geophys Res 110:D10108. https://doi.org/10.1029/2004JD005504

Birner T, Bönisch H (2011) Residual circulation trajectories and transit times into the extratropical lowermost stratosphere. Atmos Chem Phys 11(2):817–827. https://doi.org/10.5194/acp-11-817-2011

Boer G, McFarlane N, Laprise R, Henderson J, Blanchet JP (1984) The canadian climate centre spectral atmospheric general circulation model. Atmos Ocean 22(4):397–429. https://doi.org/10.1080/07055900.1984.9649208

Brewer AW (1949) Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere. Q J R Meteorol Soc 75(326):351–363. https://doi.org/10.1002/qj.49707532603

Butchart N (2014) The brewer-dobson circulation. Rev Geophys 52(2):157–184. https://doi.org/10.1002/2013RG000448

Butchart N, Cionni I, Eyring V, Shepherd T, Waugh D, Akiyoshi H, Austin J, Brühl C, Chipperfield M, Cordero E et al (2010) Chemistry-climate model simulations of twenty-first century stratospheric climate and circulation changes. J Clim 23(20):5349–5374. https://doi.org/10.1175/2010JCLI3404.1

Butchart N, Charlton-Perez AJ, Cionni I, Hardiman SC, Haynes PH, Krüger K, Kushner PJ, Newman PA, Osprey SM, Perlwitz J, Sigmond M, Wang L, Akiyoshi H, Austin J, Bekki S, Baumgaertner A, Braesicke P, Brühl C, Chipperfield M, Dameris M, Dhomse S, Eyring V, Garcia R, Garny H, Jöckel P, Lamarque JF, Marchand M, Michou M, Morgenstern O, Nakamura T, Pawson S, Plummer D, Pyle J, Rozanov E, Scinocca J, Shepherd TG, Shibata K, Smale D, Teyssèdre H, Tian W, Waugh D, Yamashita Y (2011) Multimodel climate and variability of the stratosphere. Journal of Geophysical Research: Atmospheres 116(D5):2156–2202. https://doi.org/10.1029/2010JD014995 d05102

Charney JG, Drazin PG (1961) Propagation of planetary-scale disturbances from the lower into the upper atmosphere. J Geophys Res 66(1):83–109. https://doi.org/10.1029/JZ066i001p00083

Charron M, Manzini E (2002) Gravity waves from fronts: Parameterization and middle atmosphere response in a general circulation model. J Atmos Sci 59:923–941. https://doi.org/10.1175/1520-0469(2002)059<0923:GWFFPA>2.0.CO;2

Cohen NY, Gerber EP, Bühler O (2013) Compensation between resolved and unresolved wave driving in the stratosphere: Implications for downward control. J Atmos Sci 70(12):3780–3798. https://doi.org/10.1175/JAS-D-12-0346.1

Cohen NY, Gerber EP, Bühler O (2014) What drives the brewer-dobson circulation? J Atmos Sci 71(10):3837–3855. https://doi.org/10.1175/JAS-D-14-0021.1

Collins WJ, Bellouin N, Doutriaux-Boucher M, Gedney N, Halloran P, Hinton T, Hughes J, Jones CD, Joshi M, Liddicoat S, Martin G, O'Connor F, Rae J, Senior C, Sitch S, Totterdell I, Wiltshire A, Woodward S (2011) Development and evaluation of an earth-system model—HadGEM2. Geosci Model Dev 4:1051–1075. https://doi.org/10.5194/gmd-4-1051-2011

Dietmüller S, Garny H, Plöger F, Jöckel P, Duy C (2017) Effects of mixing on resolved and unresolved scales on stratospheric age of air. Atmos Chem Phys 17:7703–7719. https://doi.org/10.5194/acp-17-7703-2017

Dietmüller S, Eichinger R, Garny H, Birner T, Boenisch H, Pitari G, Mancini E, Visioni D, Stenke A, Revell L, Rozanov E, Plummer DA, Scinocca J, Jöckel P, Oman L, Deushi M, Kiyotaka S, Kinnison DE, Garcia R, Morgenstern O, Zeng G, Stone KA, Schofield R (2018) Quantifying the effect of mixing on the mean age of air in ccmval-2 and ccmi-1 models. Atmos Chem Phys 18:6699–6720. https://doi.org/10.5194/acp-18-6699-2018

Dobson GMB (1956) Origin and distribution of the polyatomic molecules in the atmosphere. Proc R Soc Lond A 236(1205):187–193. https://doi.org/10.1098/rspa.1956.0127

Dobson GMB, Harrison DN, Lawrence J (1929) Measurements of the amount of ozone in the Earth's atmosphere and its relation to other geophysical conditions—Part III. In: Proceedings of the royal society of London A: mathematical, physical and engineering sciences 122(790):456–486, https://doi.org/10.1098/rspa.1929.0034

Edmon HR, Hoskins BJ, McIntyre ME (1980) Eliassen–Palm cross sections for the troposphere. J Atmos Sci 37(12):2600–2616. https://doi.org/10.1175/1520-0469(1980)037<2600:EPCSFT>2.0.CO;2

Eichinger R, Jöckel P (2014) The generic MESSy submodel TENDENCY (v1.0) for process-based analyses in earth system models. Geosci Model Dev 7(4):1573–1582. https://doi.org/10.5194/gmd-7-1573-2014

Eichinger R, Dietmüller S, Garny H, Šácha P, Birner T, Böhnisch H, Pitari G, Visioni D, Stenke A, Rozanov E, Revell L, Plummer DA, Jöckel P, Oman L, Deushi M, Kinnison DE, Garcia R, Morgenstern O, Zeng G, Stone KA, Schofield R (2019) The influence of mixing on the stratospheric age of air changes in the 21st century. Atmos Chem Phys 19:921–940. https://doi.org/10.5194/acp-19-921-2019

Eliassen A, Palm E (1960) On the transfer of energy in stationary mountain waves. Geofisica Int 22(3):1

Flury T, Wu D, Read W (2013) Variability in the speed of the brewer-dobson circulation as observed by aura/mls. Atmos Chem Phys 13:4563–4575. https://doi.org/10.5194/acp-13-4563-2013

Fomichev VI, Ward WE, Beagley SR, McLandress C, McConnell JC, McFarlane NA, Shepherd TG (2002) Extended canadian middle atmosphere model: zonal-mean climatology and physical parameterizations. J Geophys Res 107(D10):ACL 9-1–ACL 9-14. https://doi.org/10.1029/2001JD000479

Fritts DC, Alexander MJ (2003) Gravity wave dynamics and effects in the middle atmosphere. Rev Geophys 41(1):1003. https://doi.org/10.1029/2001RG000106

Garny H, Birner T, Bönisch H, Bunzel F (2014) The effects of mixing on age of air. J Geophys Res 119(12):7015–7034. https://doi.org/10.1002/2013JD021417

Geller MA, Alexander MJ, Love PT, Bacmeister J, Ern M, Hertzog A, Manzini E, Preusse P, Sato K, Scaife A, Zhou T (2013) A comparison between gravity wave momentum fluxes in observations and climate models. J Clim 26:6383–6405. https://doi.org/10.1175/JCLI-D-12-00545.1

Gerber EP (2012) Stratospheric versus tropospheric control of the strength and structure of the brewer-dobson circulation. J Atmos Sci 69(9):2857–2877. https://doi.org/10.1175/JAS-D-11-0341.1

Gerber EP, Butler A, Calvo N, Charlton-Perez A, Giorgetta M, Manzini E, Perlwitz J, Polvani LM, Sassi F, Scaife AA, Shaw TA, Son SW, Watanabe S (2012) Assessing and understanding the impact of stratospheric dynamics and variability on the earth system. Bull Am Meteorol Soc 93:845–859. https://doi.org/10.1175/BAMS-D-11-00145.1

Giorgetta MA, Manzini E, Roeckner E (2002) Forcing of the quasi-biennial oscillation from a broad spectrum of atmospheric waves. Geophys Res Lett 29(8):86-1–86-4 10.1029/2002GL014756

Gregory D, Shutts GJ, Mitchell JR (1998) A new gravity-wave-drag scheme incorporating anisotropic orography and low-level wave breaking: Impact upon the climate of the uk meteorological office unified model. Q J R Meteorol Soc 124(546):463–493. https://doi.org/10.1002/qj.49712454606

Hall TM, Plumb RA (1994) Age as a diagnostic of stratospheric transport. J Geophys Res 99(D1):1059–1070. https://doi.org/10.1029/93JD03192

Hardiman SC, Haynes PH (2008) Dynamical sensitivity of the stratospheric circulation and downward influence of upper level perturbations. J Geophys Res 113(D23):103. https://doi.org/10.1029/2008JD010168

Harnik N, Lindzen RS (2001) The effect of reflecting surfaces on the vertical structure and variability of stratospheric planetary waves. J Atmos Sci 58:2872–2894. https://doi.org/10.1175/1520-0469(2001)058<2872:TEORSO>2.0.CO;2

Haynes PH, McIntyre ME, G ST (1991) On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces. J Atmos Sci 48(4):651–678. https://doi.org/10.1175/1520-0469(1991)048

Hines CO (1997) Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere. J Atmos Solar-Terrestrial Phys 59:371–386. https://doi.org/10.1016/S1364-6826(96)00079-X

Holton JR (1982) The role of gravity wave-induced drag and diffusion in the momentum budget of the mesosphere. J Atmos Sci 39:791–799. https://doi.org/10.1175/1520-0469(1982)039<0791:TROGWI>2.0

Holton JR (1983) The influence of gravity wave breaking on the general circulation of the middle atmosphere. J Atmos Sci 40:2497–2507. https://doi.org/10.1175/1520-0469(1983)040<2497:TIOGWB>2.0.CO;2

Holton JR, Alexander MJ (2000) The role of waves in the transport circulation of the middle atmosphere. Geophys Monograph Ser 123:21–35. https://doi.org/10.1029/GM123p0021

Houghton JT (1978) The stratosphere and mesosphere. Q J R Meteorol Soc 104(439):1–29. https://doi.org/10.1002/qj.49710443902

Iwasaki T, Sumi A (1986) Impact of envelope orography on jma's hemispheric nwp forecasts for winter circulation. J Meteorol Soc Jp Ser II 64(2):245–258. https://doi.org/10.2151/jmsj1965.64.2_245

Jöckel P, Sander R, Kerkweg A, Tost H, Lelieveld J (2005) Technical note: the modular earth submodel system (MESSy)—a new approach towards earth system modelling. Atmos Chem Phys 5:433–444. https://doi.org/10.5194/acp-5-433-2005

Jöckel P, Kerkweg A, Pozzer A, Sander R, Tost H, Riede H, Baumgaertner A, Gromov S, Kern B (2010) Development cycle 2 of the modular earth submodel system (MESSy2). Geosci Model Dev 3:717–752. https://doi.org/10.5194/gmd-3-717-2010

Jöckel P, Holger TO, Pozzer A, Kunze M, Kirner O, Brenninkmeijer C, Brinkop S, Cai D, Dyroff C, Eckstein J, Frank F, Garny H, Gottschaldt KD, Graf P, Grewe V, Kerkweg A, Kern B, Matthes S, Mertens M, Meul S, Neumaier M, Nützel M, Oberländer-Hayn S,

Ruhnke R, Runde T, Sander R, Scharffe D, Zahn A (2016) Earth system chemistry integrated modelling (ESCiMO) with the modular earth submodel system (MESSy, version 2.51). Geosci Model Dev 9:1153–1200. https://doi.org/10.5194/gmd-9-1153-2016

Kidston J, Scaife AA, Hardiman SC, Mitchell DM, Butchart N, Baldwin MP, Gray LJ (2015) Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. Nat Geosci 8:433–440. https://doi.org/10.1038/ngeo2424

Kim YJ, Arakawa A (1995) Improvement of orographic gravity wave parameterization using a mesoscale gravity wave model. J Atmos Sci 52(11):1875–1902. https://doi.org/10.1175/1520-0469(1995)052<1875:IOOGWP>2.0.CO;2

Kim YJ, Eckermann SD, Chun HY (2003) An overview of the past, present and future of gravity wave drag parametrization for numerical climate and weather prediction models. Atmos Ocean 41(1):65–98. https://doi.org/10.3137/ao.410105

Lindzen RS (1981) Turbulence and stress owing to gravity wave and tidal breakdown. J Geophys Res 86:9707–9714. https://doi.org/10.1029/JC086iC10p09707

Lott F, Miller MJ (1997) A new subgrid-scale orographic drag parametrization: its formulation and testing. Q J R Meteorol Soc 123:101–127. https://doi.org/10.1002/qj.49712353704

Manzini E, McFarlane NA (1998) The effect of varying the source spectrum of a gravity wave parameterization in a middle atmosphere general circulation model. J Geophys Res 103(D24):31523–31539. https://doi.org/10.1029/98JD02274

Martin GM, Bellouin N, Collins WJ, Culverwell ID, Halloran PR, Hardiman SC, Hinton TJ, Jones CD, McDonald RE, McLaren AJ, O'Connor FM, Roberts MJ, Rodriguez JM, Woodward S, Best MJ, Brooks ME, Brown AR, Butchart N, Dearden C, Derbyshire SH, Dharssi I, Doutriaux-Boucher M, Edwards JM, Falloon PD, Gedney N, Gray LJ, Hewitt HT, Hobson M, Huddleston MR, Hughes J, Ineson S, Ingram WJ, James PM, Johns TC, Johnson CE, Jones A, Jones CP, Joshi MM, Keen AB, Liddicoat S, Lock AP, Maidens AV, Manners JC, Milton SF, Rae JGL, Ridley JK, Sellar A, Senior CA, Totterdell IJ, Verhoef A, Vidale PL, Wiltshire A (2011) The HadGEM2 family of met office unified model climate configurations. Geosci Model Dev 4(3):723–757. https://doi.org/10.5194/gmd-4-723-2011

Matsuno T (1892) A quasi one-dimensional model of the middle atmosphere circulation interacting with internal gravity waves. J Meteorol Soc Jpn 60:215–217. https://doi.org/10.2151/jmsj1965.60.1_215

Matsuno T (1970) Vertical propagation of stationary planetary waves in the winter northern hemisphere. J Atmos Sci 27:871–883. https://doi.org/10.1175/1520-0469(1970)027<0871:VPOSPW>2.0

McFarlane NA (1987) The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere. J Atmos Sci 44(14):1775–1800. https://doi.org/10.1175/1520-0469(1987)044<1775:TEOOEG>2.0.CO;2

Medvedev AS, Klaassen GP (1995) Vertical evolution of gravity wave spectra and the parameterization of associated wave drag. J Geophys Res 100(D12):25841–25853. https://doi.org/10.1029/95JD02533

Morgenstern O, Hegglin MI, Rozanov E, O'Connor FM, Abraham NL, Akkyoshi H, Archibald AT, Bekki S, Butchart N, Chipperfield MP, Deushi M, Dhomse SS, Garcia RR, Hardiman SC, Horowitz LW, Jöckel P, Josse B, Kinnison D, Lin MY, Mancini E, Manyin ME, Marchand M, Marecal V, Michou M, Oman LD, Pitari G, Plummer DA, Revell LE, Saint-Martin D, Schofield R, Stenke A, Stone K, Sudo K, Tanaka TY, Tilmes S, Yamashita Y, Yoshida K, Zeng G (2017) Review of the global models used within the chemistry-climate model initiative (ccmi). Geosci Model Dev 10:639–671. https://doi.org/10.5194/gmd-10-639-2017

Neu JL, Plumb RA (1999) Age of air in a leaky pipe model of stratospheric transport. J Geophys Res 104(D16):19243–19255. https://doi.org/10.1029/1999JD900251

Oort AH, Yienger JJ (1996) Observed interannual variability in the hadley circulation and its connection to enso. J Clim 9(11):2751–2767. https://doi.org/10.1175/1520-0442(1996)009<2751:OIVITH>2.0.CO;2

Palmer TN, Mansfield DA (1986) A study of wintertime circulation anomalies during past el niño events using a high resolution general circulation model. i: Influence of model climatology. Q J R Meteorol Soc 112(473):613–638. https://doi.org/10.1002/qj.49711247304

Palmer TN, Shutts GJ, Swinbank R (1986) Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parametrization. Q J R Meteorol Soc 112:1001–1039. https://doi.org/10.1002/qj.49711247406

Preusse P, Eckermann SD, Ern M, Oberheide J, Picard RH, Roble RG, Riese M, Russell JM (2009) Global ray tracing simulations of the saber gravity wave climatology. J Geophys Res 114:D08126. https://doi.org/10.1029/2008JD011214

Randel WJ, Garcia R, Wu F (2008) Dynamical balances and tropical stratospheric upwelling. J Atmos Sci 65(11):3584–3595. https://doi.org/10.1175/2008JAS2756.1

Richter JH, Sassi F, Garcia RR (2010) Toward a physically based gravity wave source parameterization in a general circulation model. J Atmos Sci 67(1):136–156. https://doi.org/10.1175/2009JAS3112.1

Šácha P, Lilienthal F, Jacobi C, Pišoft P (2016) Influence of the spatial distribution of gravity wave activity on the middle atmospheric dynamics. Atmos Chem Phys 16:15755–15775. https://doi.org/10.5194/acp-16-15755-2016

Šácha P, Miksovsky J, Pišoft P (2018) Interannual variability in the gravity wave drag—vertical coupling and possible climate links. Earth Syst Dyn 9:647–661. https://doi.org/10.5194/esd-9-647-2018

Samtleben N, Jacobi C, Pišoft P, Šácha P, Kuchař (2019) Effect of latitudinally displaced gravity wave forcing in the lower stratosphere on the polar vortex stability. Ann Geophys Discuss p Rev. https://doi.org/10.5194/angeo-2019-15

Sato K, Hirano S (2019) The climatology of the brewer-dobson circulation and the contribution of gravity waves. Atmos Chem Phys 19(7):4517–4539. https://doi.org/10.5194/acp-19-4517-2019

Sato K, Watanabe S, Kawatani Y, Tomikawa Y, Miyazaki K, Takahashi M (2009) On the origins of mesospheric gravity waves. Geophys Res Lett 36:L19801. https://doi.org/10.1029/2009GL039908

Sato K, Tateno S, Watanabe S, Kawatani Y (2012) Gravity wave characteristics in the southern hemisphere revealed by a high-resolution middle-atmosphere general circulation model. J Atmos Sci 69:1378–1396. https://doi.org/10.1175/JAS-D-11-0101.1

Scaife AA, Butchart N, Warner CD, Stainforth D, Norton W, Austin J (2000) Realistic quasi-biennial oscillations in a simulation of the global climate. Geophys Res Lett 27:3481–3484. https://doi.org/10.1029/2000GL011625

Scinocca JF (2003) An accurate spectral nonorographic gravity wave drag parameterization for general circulation models. J Atmos Sci 60(4):667–682. https://doi.org/10.1175/1520-0469(2003)060<0667:AASNGW>2.0.CO;2

Scinocca JF, McFarlane NA (2000) The parametrization of drag induced by stratified flow over anisotropic orography. Q J R Meteorol Soc 126(568):2353–2393. https://doi.org/10.1002/qj.49712656802

Shepherd TG (2007) Transport in the middle atmosphere. J Meteorol Soc Jpn 85:165–191. https://doi.org/10.2151/jmsj.85B.165

Sigmond M, Shepherd TG (2014) Compensation between resolved wave driving and parameterized orographic gravity wave driving

of the brewer-dobson circulation and its response to climate change. J Clim 27(14):5601–5610. https://doi.org/10.1175/JCLI-D-13-00644.1

Simpson IR, Blackburn M, Haigh JD (2009) The role of eddies in driving the tropospheric response to stratospheric heating perturbations. J Atmos Sci 66(5):1347–1365. https://doi.org/10.1175/2008JAS2758.1

Solomon S, Rosenlof K, Portmann R, Daniel J, Davis S, Sanford T, Plattner GK (2010) Contributions of stratospheric water vapor to decadal changes in the rate of global warming. Science 327(5970):1219–1223. https://doi.org/10.1126/science.1182488

Song IS, Chun HY (2005) Momentum flux spectrum of convectively forced internal gravity waves and its application to gravity wave drag parameterization. J Atmos Sci 62:107–124. https://doi.org/10.1175/JAS-3363.1

Thompson DWJ, Solomon S, Kushner PJ, England PJ, Grise KM, Karoly DJ (2011) Signatures of the antarctic ozone hole in southern hemisphere surface climate change. Nat Geosci 4:741–749. https://doi.org/10.1038/ngeo1296

Tibaldi S (1986) Envelope orography and maintenance of the quasi-stationary circulation in the ecmwf global models. In: Saltzman B, Benzi R, Wiin-Nielsen AC (eds) Anomalous atmospheric flows and blocking, advances in geophysics, vol 29. Elsevier, Amsterdam, pp 339–374. https://doi.org/10.1016/S0065-2687(08)60045-X

Wallace JM, Tibaldi S, Simmons AJ (1983) Reduction of systematic forecast errors in the ecmwf model through the introduction of an envelope orography. Q J R Meteorol Soc 109(462):683–717. https://doi.org/10.1002/qj.49710946202

Warner CD, McIntyre ME (1996) On the propagation and dissipation of gravity wave spectra through a realistic middle atmosphere. J Atmos Sci 53(22):3213–3235. https://doi.org/10.1175/1520-0469(1996)053<3213:OTPADO>2.0.CO;2

Waugh D, Hall T (2002) Age of stratospheric air: Theory, observations, and models. Rev Geophys 40(4):1-1–1-26. https://doi.org/10.1029/2000RG000101